Stochastic Processes and their Applications 36 (1990) 1-14
North-Holland

LARGE DEVIATIONS AND THE INTERNAL FLUCTUATIONS OF CRITICAL MEAN FIELD SYSTEMS

F. PAPANGELOU

Department of Mathematics, The University of Manchester, Manchester M13 9PL, UK

Received 22 June 1989

It is well known that the fluctuations of critical mean field systems are non-Gaussian. However, as shown in Papangelou (1989), the internal or second order fluctuations of critical Curie-Weiss models, i.e. fluctuations between subsystems, are Gaussian. In the present article we prove the same result by a large deviations approach which applies to more general mean field systems.

critical mean field fluctuations * large deviations * entropy * Laplace's method * Brownian bridge

1. Introduction

Systems of dependent random variables ('spins') $X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)}$ with joint distributions of the form

$$P(X_1^{(n)} \in dx_1, X_2^{(n)} \in dx_2, \ldots, X_n^{(n)} \in dx_n)$$
$$= \text{const.} \exp\left\{\beta n G\left(\frac{x_1+\cdots+x_n}{n}\right)\right\} \prod_{i=1}^{n} \rho(dx_i)$$

have been considered extensively in the literature because of their significance for mean field models in statistical mechanics. One of the most interesting examples, the basic Curie-Weiss model, has $G(x)=\frac{1}{2}x^2$ and $\rho(\{1\})=\rho(\{-1\})=\frac{1}{2}$ [8, 15]. This and other examples exhibit 'phase transition' at a critical value β_c of β: for $\beta<\beta_c$ the asymptotic distribution of $S_n/\sqrt{n}$ as $n\to\infty$, where S_n is the total field $X_1^{(n)}+\cdots+X_n^{(n)}$, is Gaussian, while for $\beta=\beta_c$ a more severely scaled sum such as $S_n/n^{3/4}$, say, has a non-Gaussian asymptotic distribution; for $\beta>\beta_c$ the weak law of large numbers fails. (We assume for convenience that the mean of the probability measure ρ is 0.) The existence of a non-Gaussian limit in the critical case was first proved in [14] for the basic Curie-Weiss model and since then there have been many studies of this phenomenon, which is in fact a large deviations phenomenon, implicitly treated as such in [14], and explicitly so in [6].

This result on critical behaviour has been generalised in various directions [7, 11, 3, 4, 2] and has also been extended to mean field systems undergoing time evolution [5, 2]. Implicit in the treatment of the case $G(x)=\frac{1}{2}x^2$ in [7] is the fact that $\exp\{\frac{1}{2}x^2\}$ is the moment generating function of the standard normal distribution.

0304-4149/90/$03.50

This was already exploited by Kac in [10] and leads in fact to a de Finetti representation of the joint distribution of $X_1^{(n)}, \ldots, X_n^{(n)}$ as a mixture of i.i.d.'s. (This line of thought was pursued in [4], where the above result was extended to sequences of moment generating functions $\exp G_n(x)$.) However, the result is independent of this fact and, for instance, [11] and [3] rely exclusively on large deviations and Laplace's method.

For the case $G(x)=\frac{1}{2}x^2$ we showed in [12] that, despite the non-Gaussian nature of the critical fluctuations, the fluctuations between sub-blocks of spins are Gaussian. If, for example, the system consists of $2n$ random variables $X_1^{(2n)}, \ldots, X_{2n}^{(2n)}$, then the asymptotic distribution of

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i^{(2n)} - \frac{1}{\sqrt{n}}\sum_{i=n+1}^{2n} X_i^{(2n)}$$

is Gaussian. With this as a starting point we proved in [12] a number of limit theorems, of which the following is perhaps the most illuminating representative.

Define the standard polygonal line

$$\xi_n(t) = X_1^{(n)} + \cdots + X_{[nt]}^{(n)} + (nt-[nt])X_{[nt]+1}^{(n)}, \quad 0 \leq t \leq 1.$$

Theorem 1 [12]. *Let $G(x)=\frac{1}{2}x^2$, $\beta=1$ and suppose that the probability measure ρ has mean* 0 *and variance* 1, *and that*

$$\psi(s) := \log \int_{-\infty}^{\infty} e^{sx}\rho(dx) < \tfrac{1}{2}s^2$$

for all $s \neq 0$. Then there is (an even integer) $p>2$ such that

$$\frac{\xi_n(t)}{n^{1-1/p}} = \frac{S_n}{n^{1-1/p}}\, t + \frac{1}{n^{1/2-1/p}}\, \eta_n(t), \quad 0 \leq t \leq 1,$$

where as $n \to \infty$ the random variable $S_n/n^{1-1/p}$ and the process $\eta_n(t)$, $0 \leq t \leq 1$, converge jointly in distribution, the former to a random variable Y with probability density function const. $\exp(-\lambda y^p)$, $-\infty < y < \infty$, *and the latter to a Brownian bridge independent of Y.* □

Another theorem, which we will not repeat here, states that a differently scaled polygonal line is, asymptotically, the sum $Yt+\zeta(t)$ of a random straight line Yt as in Theorem 1 and an independent standard Brownian motion $\zeta(t)$. We refer the reader to [12] for further discussion of these results.

The proofs given in [12] relied on the de Finetti representation mentioned above. In the present paper we give, as promised in [12], an alternative approach which relies exclusively on large deviations techniques and requires no assumption that $\exp G(x)$ is a moment generating function. We will not re-establish here analogues of all the limit theorems stated in [12]; we will only prove the convergence of finite-dimensional distributions behind the assertion of Theorem 1. The result obtained can be described informally as follows. Suppose the system consists of kn spins grouped into k subsets (k fixed) of n spins each, and let $S_n^1, S_n^2, \ldots, S_n^k$ be the block spins of these subsets (S_n^j is $\sum_i X_i^{(nk)}$ with the summation restricted to the jth subset). Under the hypotheses set out below there is $p > 2$ such that, if the linear space $\mathbb{R}^k$ is scaled by the factor $n^{-(1-1/p)}$ in the direction of the principal diagonal vector $(1, 1, \ldots, 1)$ and by the factor $n^{-1/2}$ in all directions orthogonal to the diagonal, then the distribution of the random vector $(S_n^1, S_n^2, \ldots, S_n^k)$ in $\mathbb{R}^k$ converges, as $n \to \infty$, to the product of a non-Gaussian distribution along the principal diagonal and a standard rotationally invariant Gaussian distribution on the $(k-1)$-dimensional linear subspace orthogonal to the diagonal.

The proofs given below require nothing more than Cramér's original one-dimensional large deviations estimates for tail probabilities: no local large deviations theorem will be used. Prior to proving the main result (Theorems 3 and 4) we tread in Section 2 on the familiar territory of the Laplace approach (cf. [11, 3]), albeit along a somewhat novel path, culminating in the statement of the non-Gaussian fluctuations of the field. We have two purposes in doing this. The first is to state and prove on the way the new Proposition 3 and its corollary, which are essential for Section 3. The second is to show how all proofs in Sections 2 and 3 can be based on Proposition 1 and the quite weak assumptions made in Section 2. These assumptions cover cases where $G(x)$ is genuinely quadratic throughout its domain.

I take this opportunity to thank Gérard Ben Arous for making available to me a preprint of [2] and for a useful conversation on the subject.

2. The fluctuations of the field

Throughout the paper we assume that ρ is a probability measure on the Borel σ-field of the real line, with mean 0 and variance 1 and such that $\int_{-\infty}^{\infty} e^{sx}\rho(dx) < \infty$ for all $s \in (-\infty, \infty)$. The cumulant generating function of ρ,

$$\psi(s) = \log \int_{-\infty}^{\infty} e^{sx}\rho(dx)$$

is then convex and real analytic on $(-\infty, \infty)$ and satisfies $\psi'(0) = 0$, $\psi''(0) = 1$. The entropy function of ρ is

$$H(x) = \sup_t \{tx - \psi(t)\}.$$

If ζ denotes the supremum and θ the infimum of the support of ρ ($-\infty \leq \theta < 0 < \zeta \leq \infty$), then H is convex and real analytic in (θ, ζ) and

$$H(x) = \tfrac{1}{2}x^2 + O(x^3) \quad \text{as } x \to 0.$$

The derivative ψ' is a strictly increasing mapping of $(-\infty, \infty)$ on to (θ, ζ) and if, given $x \in (\theta, \zeta)$, we let s be the unique real number such that $\psi'(s) = x$, then $H(x) = s\psi'(s) - \psi(s)$. The inverse of $x = \psi'(s)$ is $s = H'(x)$ and it can be shown that

$H'(x) \to \infty$ as $x \to \zeta$ and $H(x)/x \to \infty$ as $x \to \infty$. Analogously for negative x. We refer the reader to [1] and [8] for basic facts.

All our proofs below will be based entirely on three classical facts about one-dimensional large deviations, going back to Cramér's work. We collect these facts in the following proposition.

Denote by F_n the (right-continuous) distribution function of $(1/n)\sum_{i=1}^n W_i$ where $W_1, \ldots, W_n$ are i.i.d. random variables with distribution ρ.

Proposition 1. (i) *For all* $x \geq 0$,

$$1 - F_n(x) \leq \exp\{-nH(x)\}. \tag{1}$$

(ii) *If* $0 < b < \zeta$, *then there exists a constant* $C(b)$ *such that for all* $x \in [0, b]$,

$$x(1 - F_n(x)) \leq \frac{C(b)}{\sqrt{n}} \exp\{-nH(x)\}. \tag{2}$$

(iii) *Let* $p > 2$ *and* $0 < a < b < \infty$. *If* $n \to \infty$ *then, for* $an^{-1/p} \leq x \leq bn^{-1/p}$,

$$x(1 - F_n(x)) = \frac{1 + \mathrm{o}(1)}{\sqrt{2\pi n}} \exp\{-nH(x)\} \tag{3}$$

where $\mathrm{o}(1) \to 0$ *uniformly for* x *in the above range.* □

Analogous statements can of course be made for negative x.

The first statement is an immediate and familiar consequence of the definition of H, since

$$1 - F_n(x) \leq \exp(-tnx)E(\mathrm{e}^{t\sum_{i=1}^n W_i}) = \exp\{-n(tx - \psi(t))\}$$

for arbitrary t. The second and third statements follow from the considerations of [9, pp. 548–553]. See also [13, Chapter VIII].

In what follows we will be concerned with a system of random variables $X_1^{(n)}, X_2^{(n)}, \ldots, X_n^{(n)}$ having joint distribution

$$P(X_1^{(n)} \in \mathrm{d}x_1, X_2^{(n)} \in \mathrm{d}x_2, \ldots, X_n^{(n)} \in \mathrm{d}x_n) = \mathscr{Z}_n^{-1} \exp\left\{nG\left(\frac{x_1 + \cdots + x_n}{n}\right)\right\} \prod_{i=1}^n \rho(\mathrm{d}x_i), \tag{4}$$

where $G(x)$ is defined and finite for all x and $\mathscr{Z}_n$ is the appropriate norming constant. We make the following assumptions about G.

(I) G is everywhere differentiable, its derivative G' is bounded on bounded sets and $G'(x) = x(1 + \mathrm{o}(1))$ as $x \to 0$.

(II) $G(x) < H(x)$ for all $x \neq 0$.

(III) $\limsup G(x)/H(x) < 1$ as $x \to \zeta$ or $x \to \theta$.

(IV) There exist $\lambda > 0$ and $p > 2$ such that $G(x) - H(x) = -\lambda|x|^p + \mathrm{o}(|x|^p)$ as $x \to 0$.

Note that (I) implies $G(x) = \frac{1}{2}x^2 + \mathrm{o}(x^2)$ as $x \to 0$ and that (III) is trivially true if $\zeta < \infty$ and $\theta > -\infty$. Condition (III) is sufficient to guarantee that

$\int_{-\infty}^{\infty} \exp G(x)\rho(\mathrm{d}x) < \infty$ and hence that $\mathscr{Z}_n < \infty$ for all n, even if $\zeta = \infty$ or $\theta = \infty$, as will be shown below. It is also less restrictive than the condition "$|G(x)| \leq a|x| + b$ for some a, b and all sufficiently large $|x|$", assumed in [11] and [3], since the latter excludes the function $G(x) = \frac{1}{2}x^2$ if $\zeta = \infty$ or $\theta = -\infty$.

Lemma 1. *In order that* (I), (II) *and* (III) *hold, it is sufficient that* G' *exist,*

$$0 < G'(\psi'(s))/s < 1 \quad \textit{for } s \neq 0$$

and

$$\limsup G'(\psi'(s))/s < 1 \quad \textit{as } s \to \pm\infty. \qquad \square$$

In fact the given conditions imply, for instance, that $G'(x) < H'(x)$ for all $x \in (0, \zeta)$ and that there exist $\delta > 0$ and $x_0 \in (0, \zeta)$ such that $G'(x) < (1-\delta)H'(x)$ for $x \in (x_0, \zeta)$. It follows that $G(x) - H(x)$ is decreasing for $x > 0$ and that $G(x) \leq (1-\delta)H(x) + G(x_0) - (1-\delta)H(x_0)$ for $x \geq x_0$. The latter implies (III) if $\zeta = \infty$, since $H(x)/x \to \infty$ as $x \to \infty$.

As an example, if $G(x) = \frac{1}{2}x^2$ and if $\psi'''(s) < 0$ for $s > 0$ and $\psi'''(s) > 0$ for $s < 0$ (GHS inequality, cf. [7, p. 137]), then (I), (II), (III) hold.

If $S_n = X_1^{(n)} + X_2^{(n)} + \cdots + X_n^{(n)}$, then the distribution of S_n/n is given by

$$P(\alpha < S_n/n \leq \beta) = \frac{1}{\mathscr{Z}_n} \int_\alpha^\beta \exp\{nG(x)\}\, \mathrm{d}F_n(x),$$

where we use the convention that $\int_\alpha^\beta$ denotes $\int_{(\alpha,\beta]}$. If $0 \leq \alpha < \beta < \infty$, an integration by parts shows that

$$\begin{aligned} &\textstyle\int_\alpha^\beta \exp\{nG(x)\}\, \mathrm{d}F_n(x) \\ &\quad = n \int_\alpha^\beta \exp\{nG(x)\} G'(x)(1 - F_n(x))\, \mathrm{d}x \\ &\qquad + (1 - F_n(\alpha)) \exp\{nG(\alpha)\} - (1 - F_n(\beta)) \exp\{nG(\beta)\}, \end{aligned} \tag{5}$$

where by (1) and (II) each of the last two terms is bounded by 1.

The propositions which follow are formulated for positive a, b but it should be obvious that analogous statements hold for negative values.

Proposition 2. *For* $0 \leq a < b < \infty$,

$$\lim_{n\to\infty} n^{-1/2+1/p} \int_{an^{-1/p}}^{bn^{-1/p}} \exp\{nG(x)\}\, \mathrm{d}F_n(x) = \frac{1}{\sqrt{2\pi}} \int_a^b \exp(-\lambda z^p)\, \mathrm{d}z.$$

Proof. By (5), (1) and (II),

$$\begin{aligned} &\int_{an^{-1/p}}^{bn^{-1/p}} \exp\{nG(x)\}\, \mathrm{d}F_n(x) \\ &\quad = n \int_{an^{-1/p}}^{bn^{-1/p}} \exp\{nG(x)\} G'(x)(1 - F_n(x))\, \mathrm{d}x + \mathrm{O}(1). \end{aligned}$$

If $a>0$, then by (2) the last term is in fact o(1) and (I), (3) and (IV) imply that the right-hand side is equal to

$$\frac{1+\mathrm{o}(1)}{\sqrt{2\pi}}\,n^{1/2}\int_{an^{-1/p}}^{bn^{-1/p}} \exp\{n(G(x)-H(x))\}\,\mathrm{d}x+\mathrm{o}(1)$$

$$=\frac{1+\mathrm{o}(1)}{\sqrt{2\pi}}\,n^{1/2}\int_{an^{-1/p}}^{bn^{-1/p}} \exp\{n(-\lambda x^p+\mathrm{o}(x^p))\}\,\mathrm{d}x+\mathrm{o}(1)$$

$$=\frac{1+\mathrm{o}(1)}{\sqrt{2\pi}}\,n^{1/2-1/p}\int_{a}^{b} \exp\left\{-\lambda z^p+n\mathrm{o}\left(\frac{z^p}{n}\right)\right\}\mathrm{d}z+\mathrm{o}(1)$$

which proves the result for $a>0$. Now notice that by (I), (2) and (II),

$$n\int_0^{an^{-1/p}} \exp\{nG(x)\}G'(x)(1-F_n(x))\,\mathrm{d}x$$

$$\leqslant Cn^{1/2}\int_0^{an^{-1/p}} \exp\{n(G(x)-H(x))\}\,\mathrm{d}x\leqslant Can^{1/2-1/p}$$

for some constant C, and by taking a sufficiently small a we can extend the result to the case where the lower limit of integration is 0. □

Proposition 3. *If $a_n\to a>0$ as $n\to\infty$, then for any finite $u<v$,*

$$\lim_{n\to\infty}\int_{a_nn^{-1/p}+un^{-1/2}}^{a_nn^{-1/p}+vn^{-1/2}} \exp\{nG(x)\}\,\mathrm{d}F_n(x)=\frac{v-u}{\sqrt{2\pi}}\exp(-\lambda a^p).$$

Proof. Just as in the proof of Proposition 2, we see that the integral on the left is equal to

$$\frac{1+\mathrm{o}(1)}{\sqrt{2\pi}}\,n^{1/2-1/p}\int_{a_n+un^{-1/2+1/p}}^{a_n+vn^{-1/2+1/p}} \exp\left\{-\lambda z^p+n\mathrm{o}\left(\frac{z^p}{n}\right)\right\}\mathrm{d}z+\mathrm{o}(1)$$

$$=(1+\mathrm{o}(1))\frac{v-u}{\sqrt{2\pi}}[(v-u)n^{-1/2+1/p}]^{-1}\int_{a_n+un^{-1/2+1/p}}^{a_n+vn^{-1/2+1/p}} \exp(-\lambda z^p)\,\mathrm{d}z+\mathrm{o}(1)$$

from which the result follows immediately. □

We will find it convenient to state Proposition 3 in slightly different notation. First notice that an integral such as $\int_\alpha^\beta \exp\{nG(x)\}\,\mathrm{d}F_n(x)$ can also be written as

$$\int_{\alpha\sqrt{n}}^{\beta\sqrt{n}} \exp\left\{nG\left(\frac{x}{\sqrt{n}}\right)\right\}\mathrm{d}F_n\left(\frac{x}{\sqrt{n}}\right),$$

where $F_n(x/\sqrt{n})$ as a function of x is the distribution function of $(1/\sqrt{n})\sum_{i=1}^n W_i$. (See the paragraph preceding Proposition 1.) Proposition 3 then asserts that

$$\lim_{n\to\infty}\int_{a_nn^{1/2-1/p}+u}^{a_nn^{1/2-1/p}+v} \exp\left\{nG\left(\frac{x}{\sqrt{n}}\right)\right\}\mathrm{d}F_n\left(\frac{x}{\sqrt{n}}\right)=\frac{1}{\sqrt{2\pi}}\exp(-\lambda a^p)(v-u).$$

If we define a sequence of measures μ_n, $n=1,2,\ldots$, on the Borel σ-field $\mathscr{B}$ of $\mathbb{R}$ by setting for $B \in \mathscr{B}$,

$$\mu_n(B) = \int_B \exp\left\{nG\left(\frac{x}{\sqrt{n}}\right)\right\} \mathrm{d}F_n\left(\frac{x}{\sqrt{n}}\right), \tag{6}$$

we see that Proposition 3 asserts the convergence of the measures $\mu_n(a_n n^{1/2-1/p}+B)$ (where $x+B$ is the obvious translate of B) to a constant multiple of Lebesgue measure. Denoting measures by their differentials, we can state this as follows.

Corollary 1. *If $a_n \to a > 0$ then, on any bounded interval, the sequence of measures*

$$\mu_n(a_n n^{1/2-1/p} + \mathrm{d}t)$$

converges weakly to the measure

$$\frac{1}{\sqrt{2\pi}} \exp(-\lambda a^p)\, \mathrm{d}t. \qquad \square$$

This can also be stated as follows: if $a_n \to a > 0$, then the measure $\mathscr{Z}_n P((S_n/\sqrt{n} - a_n n^{1/2-1/p}) \in \mathrm{d}t)$ converges weakly to $(1/\sqrt{2\pi})\exp(-\lambda a^p)\,\mathrm{d}t$ on every bounded interval.

Next we turn to the tails of our integrals.

Proposition 4. *There is a constant K such that for all $b > 0$,*

$$\limsup_{n\to\infty} n^{-1/2+1/p} \int_{bn^{-1/p}}^{\infty} \exp\{nG(x)\}\, \mathrm{d}F_n(x) \leq K \int_b^{\infty} \exp(-\tfrac{1}{2}\lambda z^p)\, \mathrm{d}z.$$

Proof. Choose $\varepsilon > 0$ so small that $G(x) - H(x) < -\frac{1}{2}\lambda x^p$ for $0 \leq x \leq \varepsilon$ and split the integral on the left into

$$\int_{bn^{-1/p}}^{\varepsilon} + \int_{\varepsilon}^{\zeta} := I_n + J_n.$$

Consider I_n first. By (5) it is sufficient to consider

$$I_n' = n \int_{bn^{-1/p}}^{\varepsilon} \exp\{nG(x)\} G'(x)(1 - F_n(x))\, \mathrm{d}x.$$

By (I) and (2) there is a constant K such that

$$\begin{aligned} |I_n'| &\leq Kn^{1/2} \int_{bn^{-1/p}}^{\varepsilon} \exp\{n(G(x) - H(x))\}\, \mathrm{d}x \\ &\leq Kn^{1/2} \int_{bn^{-1/p}}^{\varepsilon} \exp\{-\tfrac{1}{2}n\lambda x^p\}\, \mathrm{d}x \\ &\leq Kn^{1/2-1/p} \int_b^{\varepsilon n^{1/p}} \exp(-\tfrac{1}{2}\lambda z^p)\, \mathrm{d}z. \end{aligned} \tag{7}$$

Next consider J_n. By (III) there is $\delta > 0$ such that

$$J_n \leq \int_\varepsilon^\zeta \exp\{n(1-\delta)H(x)\}\,\mathrm{d}F_n(x)$$

$$= n(1-\delta)\int_\varepsilon^\zeta \exp\{n(1-\delta)H(x)\}H'(x)(1-F_n(x))\,\mathrm{d}x$$

$$+(1-F_n(\varepsilon))\exp\{n(1-\delta)H(\varepsilon)\}.$$

By (1) the second term on the right is at most $\exp\{-n\delta H(\varepsilon)\}$ while the first term on the right is at most

$$n(1-\delta)\int_\varepsilon^\zeta \exp\{-n\delta H(x)\}H'(x)\,\mathrm{d}x = \frac{1-\delta}{\delta}[\mathrm{e}^{-n\delta H(\varepsilon)} - \mathrm{e}^{-n\delta H(\zeta)}].$$

This and (7) imply the proposition. □

We can now deduce that the familiar statement on the non-Gaussian fluctuations of the field, referred to in the introduction, can be made under the relatively weak assumptions of this section.

Theorem 2.

(i) $$\lim_{n\to\infty} \mathcal{Z}_n / n^{1/2-1/p} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp(-\lambda|x|^p)\,\mathrm{d}x.$$

(ii) *For* $-\infty \leq a < b \leq \infty$,

$$\lim_{n\to\infty} P(a < S_n/n^{1-1/p} \leq b) = \int_a^b \exp(-\lambda|x|^p)\,\mathrm{d}x \Big/ \int_{-\infty}^{\infty} \exp(-\lambda|x|^p)\,\mathrm{d}x.$$

□

Both parts follow from Propositions 2 and 4 and the fact that

$$P(a < S_n/n^{1-1/p} \leq b) = P(an^{-1/p} < S_n/n \leq bn^{-1/p})$$

$$= \frac{1}{\mathcal{Z}_n}\int_{an^{-1/p}}^{bn^{-1/p}} \exp\{nG(x)\}\,\mathrm{d}F_n(x). \tag{8}$$

If we denote by $Q_n(\mathrm{d}x)$ the distribution of $S_n/n^{1-1/p}$,

$$Q_n(\mathrm{d}x) = P(S_n/n^{1-1/p} \in \mathrm{d}x), \tag{9}$$

then the second part of the theorem may be formulated as follows.

Corollary 2. *As* $n \to \infty$, *the sequence of distributions* $Q_n(\mathrm{d}x)$ $n = 1, 2, \ldots,$ *converges weakly to the distribution* const. $\exp(-\lambda|x|^p)\,\mathrm{d}x$. □

3. The internal fluctuations

Let k be a fixed positive integer and consider kn random variables $X_1^{(kn)}, X_2^{(kn)}, \ldots, X_{kn}^{(kn)}$ with joint distribution given by (4) but with n replaced by

kn. Define

$$S_n^j = X_{(j-1)n+1}^{(kn)} + \cdots + X_{jn}^{(kn)}, \quad j = 1, 2, \ldots, k.$$

The joint distribution of $S_n^1, S_n^2, \ldots, S_n^k$ is given by

$$P\left(\frac{S_n^1}{\sqrt{n}} \in \mathrm{d}z_1, \frac{S_n^2}{\sqrt{n}} \in \mathrm{d}z_2, \ldots, \frac{S_n^k}{\sqrt{n}} \in \mathrm{d}z_k\right)$$

$$= \frac{1}{\mathscr{Z}_{kn}} \exp\left\{knG\left(\frac{z_1 + \cdots + z_k}{k\sqrt{n}}\right)\right\} \prod_{j=1}^{k} \mathrm{d}F_n\left(\frac{z_j}{\sqrt{n}}\right).$$

In the present section we are interested in the joint distribution of

$$\frac{S_n^k}{n^{1-1/p}}; \quad \frac{S_n^1}{\sqrt{n}} - \frac{S_n^k}{\sqrt{n}}, \frac{S_n^2}{\sqrt{n}} - \frac{S_n^k}{\sqrt{n}}, \ldots, \frac{S_n^{k-1}}{\sqrt{n}} - \frac{S_n^k}{\sqrt{n}}.$$

Let $0 < a < b$, $u_1 < v_1, u_2 < v_2, \ldots, u_{k-1} < v_{k-1}$; then

$$P\left(a < \frac{S_n^k}{n^{1-1/p}} \leqslant b, u_1 < \frac{S_n^1}{\sqrt{n}} - \frac{S_n^k}{\sqrt{n}} \leqslant v_1, \ldots, u_{k-1} < \frac{S^{k-1}}{\sqrt{n}} - \frac{S^k}{\sqrt{n}} \leqslant v_{k-1}\right)$$

$$= P\left(an^{1/2-1/p} < \frac{S_n^k}{\sqrt{n}} \leqslant bn^{1/2-1/p}, \frac{S_n^k}{\sqrt{n}} + u_1 < \frac{S_n^1}{\sqrt{n}} \leqslant \frac{S_n^k}{\sqrt{n}} + v_1,\right.$$

$$\left.\ldots, \frac{S_n^k}{\sqrt{n}} + u_{k-1} < \frac{S^{k-1}}{\sqrt{n}} \leqslant \frac{S_n^k}{\sqrt{n}} + v_{k-1}\right)$$

$$= \frac{1}{\mathscr{Z}_{kn}} \int_{an^{1/2-1/p}}^{bn^{1/2-1/p}} \left[\int_{z_k+u_1}^{z_k+v_1} \cdots \int_{z_k+u_{k-1}}^{z_k+v_{k-1}} \exp\left\{knG\left(\frac{z_1 + \cdots + z_k}{k\sqrt{n}}\right)\right\}\right.$$

$$\left.\cdot \prod_{j=1}^{k-1} \mathrm{d}F_n\left(\frac{z_j}{\sqrt{n}}\right)\right] \mathrm{d}F_n\left(\frac{z_k}{\sqrt{n}}\right).$$

In terms of the notation of (6) this is

$$= \frac{1}{\mathscr{Z}_{kn}} \int_{an^{1/2-1/p}}^{bn^{1/2-1/p}} \left[\int_{z_k+u_1}^{z_k+v_1} \cdots \int_{z_k+u_{k-1}}^{z_k+v_{k-1}} \exp\left\{knG\left(\frac{z_1 + \cdots + z_k}{k\sqrt{n}}\right)\right.\right.$$

$$\left.\left. -nG\left(\frac{z_1}{\sqrt{n}}\right) - \cdots - nG\left(\frac{z_k}{\sqrt{n}}\right)\right\} \prod_{j=1}^{k-1} \mu_n(\mathrm{d}z_j)\right]$$

$$\cdot \exp\left\{nG\left(\frac{z_k}{\sqrt{n}}\right)\right\} \mathrm{d}F_n\left(\frac{z_k}{\sqrt{n}}\right). \tag{10}$$

Keeping z_k fixed, change the variables $z_1, \ldots, z_{k-1}$ in the multiple integral in square brackets as follows:

$$z_1 = z_k + t_1, \quad z_2 = z_k + t_2 \quad, \ldots, \quad z_{k-1} = z_k + t_{k-1}.$$

The integral in square brackets in (10) is then seen to be equal to

$$\int_{u_1}^{v_1}\cdots\int_{u_{k-1}}^{v_{k-1}} \exp\left\{knG\left(\frac{z_k}{\sqrt{n}}+\frac{t_1+\cdots+t_{k-1}}{k\sqrt{n}}\right)-nG\left(\frac{z_k}{\sqrt{n}}+\frac{t_1}{\sqrt{n}}\right)\right.$$
$$\left.-\cdots-nG\left(\frac{z_k}{\sqrt{n}}+\frac{t_{k-1}}{\sqrt{n}}\right)-nG\left(\frac{z_k}{\sqrt{n}}\right)\right\}\prod_{j=1}^{k-1}\mu_n(z_k+\mathrm{d}t_j). \quad (11)$$

Now recall that the entropy function $H(\cdot)$ is real analytic in (θ,ζ) and therefore has a power series expansion $H(x)=\sum_{\nu=2}^{\infty}c_\nu x^\nu$, where $c_2=\frac{1}{2}$ of course. By assumption (IV), $G(\cdot)$ must be of the form

$$G(x)=\sum_{\nu=2}^{[p]}c_\nu x^\nu-\lambda|x|^p+\mathrm{o}(|x|^p)\quad\text{as } x\to 0. \quad (12)$$

We will use (12) to expand the function appearing inside the curly brackets in (11). A little elementary algebra shows that the sum of terms of this function arising from the term $\frac{1}{2}x^2$ in (12) is

$$-\tfrac{1}{2}\left[\left(1-\frac{1}{k}\right)\sum_{j=1}^{k-1}t_j^2-\frac{2}{k}\sum_{1\leq i<j\leq k-1}t_it_j\right]=-\tfrac{1}{2}\boldsymbol{t}\boldsymbol{M}\boldsymbol{t}', \quad (13)$$

where $\boldsymbol{M}$ is the $(k-1)\times(k-1)$ matrix

$$\boldsymbol{M}=\begin{bmatrix}1-1/k & -1/k & \cdots & -1/k\\ -1/k & 1-1/k & \cdots & -1/k\\ \vdots & & \ddots & \\ -1/k & -1/k & \cdots & 1-1/k\end{bmatrix}$$

and $\boldsymbol{t}$ is the row vector $(t_1,t_2,\ldots,t_{k-1})$. The sum of terms arising from the term $c_\nu x^\nu$ $(2<\nu\leq[p])$ in (12) is

$$knc_\nu\left(\frac{z_k}{\sqrt{n}}+\frac{t_1+\cdots+t_{k-1}}{k\sqrt{n}}\right)^\nu-nc_\nu\left(\frac{z_k}{\sqrt{n}}+\frac{t_1}{\sqrt{n}}\right)^\nu-\cdots-nc_\nu\left(\frac{z_k}{\sqrt{n}}+\frac{t_{k-1}}{\sqrt{n}}\right)^\nu-nc_\nu\left(\frac{z_k}{\sqrt{n}}\right)^\nu$$
$$=nc_\nu\left(\frac{z_k}{\sqrt{n}}\right)^\nu\left[k\left(1+\frac{t_1+\cdots+t_{k-1}}{kz_k}\right)^\nu-\left(1+\frac{t_1}{z_k}\right)^\nu-\cdots-\left(1+\frac{t_{k-1}}{z_k}\right)^\nu-1\right]. \quad (14)$$

The function in square brackets in (14) is $\mathrm{O}(1/z_k^2)$ as $z_k\to\infty$, uniformly in $t_1\in[u_1,v_1],\ldots,t_{k-1}\in[u_{k-1},v_{k-1}]$. (To see this replace $1/z_k$ by x: the resulting function $f(x)$ of x is $f(0)+f'(0)x+\mathrm{O}(x^2)=\mathrm{O}(x^2)$ as $x\to 0$. As a matter of fact $\frac{1}{2}f''(0)=\nu(\nu-1)[-\frac{1}{2}\boldsymbol{t}\boldsymbol{M}\boldsymbol{t}']$). Hence (14) is

$$nc_\nu(z_k/\sqrt{n})^\nu\mathrm{O}(1/z_k^2)=\mathrm{O}(z_k^{\nu-2}n^{1-\nu/2}).$$

The same applies if $\nu=p$, i.e. to terms arising from $-\lambda|x|^p$. Finally the sum of terms that are $\mathrm{o}(|\cdot|^p)$ is $n\mathrm{o}((z_k/\sqrt{n})^p)=\mathrm{o}(z_k^pn^{1-p/2})$. Putting these facts together and changing the variable in the outer integral of (10) by setting

$$z_k=yn^{1/2-1/p}$$

we see that (10) is equal to

$$\frac{\mathscr{X}_n}{\mathscr{X}_{kn}}\int_a^b\left[\int_{u_1}^{v_1}\cdots\int_{u_{k-1}}^{v_{k-1}}\exp\{-\tfrac{1}{2}\boldsymbol{tMt}'+B_n(y,t)\}\cdot\prod_{j=1}^{k-1}\mu_n(yn^{1/2-1/p}+\mathrm{d}t_j)\right]Q_n(\mathrm{d}y) \tag{15}$$

in the notation of (8), (9), where $B_n(y,t)$ is a sum of terms over ν, $2<\nu\leqslant p$, which are of the form

$$\mathrm{O}(y^{\nu-2}n^{(1/2-1/p)(\nu-2)+1-\nu/2})=\mathrm{O}(y^{\nu-2}n^{(2-\nu)/p})$$

if $2<\nu<p$, or

$$\mathrm{o}(y^p n^{(1/2-1/p)+1-p/2})=\mathrm{o}(y^p)$$

if $\nu=p$. It follows that $B_n(y,t)$ converges to 0 as $n\to\infty$, uniformly in $y\in[a,b]$, $t_1\in[u_1,v_1],\ldots,t_{k-1}\in[u_{k-1},v_{k-1}]$.

The ratio

$$\mathscr{X}_n/\mathscr{X}_{kn}=(\mathscr{X}_n/n^{1/2-1/p})/(\mathscr{X}_{kn}/(kn)^{1/2-1/p})k^{1/2-1/p}$$

converges to $k^{-1/2+1/p}$. Hence, by Corollaries 1 and 2 and the Proposition in Section 2 of [12], (15) converges to

$$\begin{aligned}&\frac{1}{k^{1/2-1/p}}\int_a^b\left[\int_{u_1}^{v_1}\cdots\int_{u_{k-1}}^{v_{k-1}}\exp\{-\tfrac{1}{2}\boldsymbol{tMt}'\}\cdot\prod_{j=1}^{k-1}\frac{1}{\sqrt{2\pi}}\exp(-\lambda y^p)\,\mathrm{d}t_j\right]\frac{\exp(-\lambda y^p)\,\mathrm{d}y}{\int_{-\infty}^{\infty}\exp(-\lambda z^p)\,\mathrm{d}z}\\&=k^{1/p}\int_a^b\left[(2\pi)^{-(k-1)/2}k^{-1/2}\int_{u_1}^{v_1}\cdots\int_{u_{k-1}}^{v_{k-1}}\exp\{-\tfrac{1}{2}\boldsymbol{tMt}'\}\prod_{j=1}^{k-1}\mathrm{d}t_j\right]\cdot\frac{\exp(-k\lambda y^p)\,\mathrm{d}y}{\int_{-\infty}^{\infty}\exp(-\lambda z^p)\,\mathrm{d}z}.\end{aligned} \tag{16}$$

The $(k-1)$-dimensional distribution with probability density const. $\exp(-\frac{1}{2}\boldsymbol{tMt}')$ is the $(k-1)$-dimensional normal distribution with mean vector 0 and $(k-1)\times(k-1)$ covariance matrix

$$\boldsymbol{\Gamma}=\begin{bmatrix}2&1&\cdots&1\\1&2&&1\\\vdots&&\ddots&\\1&1&\cdots&2\end{bmatrix}$$

with $\det\boldsymbol{\Gamma}=k$. We have therefore proved the following theorem.

Theorem 3. *The asymptotic distribution of $S_n^k/n^{1-1/p}$ has probability density function* const. $\exp(-k\lambda|x|^p)$, $-\infty<x<\infty$. *The random vector*

$$\left(\frac{S_n^1}{\sqrt{n}}-\frac{S_n^k}{\sqrt{n}},\frac{S_n^2}{\sqrt{n}}-\frac{S_n^k}{\sqrt{n}},\ldots,\frac{S_n^{k-1}}{\sqrt{n}}-\frac{S_n^k}{\sqrt{n}}\right)$$

is asymptotically independent of $S_n^k/n^{1-1/p}$, *and its asymptotic distribution is the same as that of*

$$(Y_1 - Y_k, Y_2 - Y_k, \ldots, Y_{k-1} - Y_k),$$

where $Y_1, Y_2, \ldots, Y_k$ *are independent standard normal random variables.* $\square$

An immediate consequence is that $S_n^1/n^{1-1/p}, S_n^2/n^{1-1/p}, \ldots, S_n^k/n^{1-1/p}$ are asymptotically equal, in the sense that the difference of any two of them converges to 0 in probability. (See [12] for background.) Since $S_n^1 + S_n^2 + \cdots + S_n^k = S_{kn}$, we deduce that $S_n^k/n^{1-1/p}$ and $k^{-1}S_{kn}/n^{1-1/p} = k^{-1/p}S_{kn}/(kn)^{1-1/p}$ are asymptotically equal.

Corollary 3. *The random vector*

$$\left(\frac{S_n^1}{\sqrt{n}} - \frac{1}{k}\frac{S_{kn}}{\sqrt{n}}, \frac{S_n^2}{\sqrt{n}} - \frac{1}{k}\frac{S_{kn}}{\sqrt{n}}, \ldots, \frac{S_n^k}{\sqrt{n}} - \frac{1}{k}\frac{S_{kn}}{\sqrt{n}}\right)$$

is asymptotically independent of $S_{kn}/(kn)^{1-1/p}$, *and its asymptotic distribution is that of*

$$\left(Y_1 - \frac{1}{k}\sum_{i=1}^k Y_i, Y_2 - \frac{1}{k}\sum_{i=1}^k Y_i, \ldots, Y_k - \frac{1}{k}\sum_{i=1}^k Y_i\right)$$

where $Y_1, \ldots, Y_k$ *are as in Theorem* 3. $\square$

This follows from Theorem 3 and the fact that for $j = 1, 2, \ldots, k$,

$$\frac{S_n^j}{\sqrt{n}} - \frac{1}{k}\frac{S_{kn}}{\sqrt{n}} = \left(\frac{S_n^j}{\sqrt{n}} - \frac{S_n^k}{\sqrt{n}}\right) - \frac{1}{k}\sum_{i=1}^k \left(\frac{S_n^i}{\sqrt{n}} - \frac{S_n^k}{\sqrt{n}}\right).$$

This corollary establishes for general $G(\cdot)$, subject to assumptions (I) to (IV), the convergence of finite-dimensional distributions to those of a Brownian bridge asserted by Theorem 1 for the special case $G(x) = \frac{1}{2}x^2$, since

$$\frac{S_n^j}{n^{1-1/p}} = \frac{S_{kn}}{n^{1-1/p}}\frac{1}{k} + \frac{1}{n^{1/2-1/p}}\left(\frac{S_n^j}{\sqrt{n}} - \frac{1}{k}\frac{S_{kn}}{\sqrt{n}}\right).$$

The only thing left for a full analogue of Theorem 1 is to establish the tightness of the sequence of corresponding distributions in the space of continuous paths, but we will not go into this routine matter here. We will instead give a more geometric formulation of Theorem 3 in terms of the distribution of the random vector $\boldsymbol{V}_n = (S_n^1, S_n^2, \ldots, S_n^k)$ in $\mathbb{R}^k$. Set $\boldsymbol{d} = (1, 1, \ldots, 1) \in \mathbb{R}^k$ and let $\{\boldsymbol{e}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_k\}$ be an orthonormal base in the $(k-1)$-dimensional linear subspace of $\mathbb{R}^k$ orthogonal to $\boldsymbol{d}$.

Theorem 4. *Let*

$$\boldsymbol{V}_n = \frac{1}{k} S_{kn}\,\boldsymbol{d} + U_n^2\boldsymbol{e}_2 + U_n^3\boldsymbol{e}_3 + \cdots + U_n^k\boldsymbol{e}_k$$

be the representation of V_n in terms of the base $\{\boldsymbol{d}, \boldsymbol{e}_2, \boldsymbol{e}_3, \ldots, \boldsymbol{e}_k\}$ in $\mathbb{R}^k$. The asymptotic joint distribution of

$$\frac{1}{k}\frac{S_{kn}}{n^{1-1/p}}, \frac{U_n^2}{\sqrt{n}}, \frac{U_n^3}{\sqrt{n}}, \ldots, \frac{U_n^k}{\sqrt{n}},$$

as $n \to \infty$, is that of k independent random variables $Z_1, Z_2, \ldots, Z_k$, where Z_1 has probability density function const. $\exp(-k\lambda|x|^p)$, $-\infty < x < \infty$, *and $Z_2, Z_3, \ldots, Z_k$ are standard normal.* □

In the context of [3] the direction of $\boldsymbol{d}$ is the direction of degeneracy: if $\boldsymbol{V}_n$ is scaled by the factor $n^{-(1-1/p)}$, then its asymptotic distribution is carried by the one-dimensional subspace generated by $\boldsymbol{d}$. To get a genuinely k-dimensional asymptotic distribution one must scale by the factor $n^{-(1-1/p)}$ in the direction of $\boldsymbol{d}$ and by the factor $n^{-1/2}$ in all directions orthogonal to $\boldsymbol{d}$. Theorem 4 is an immediate consequence of Theorem 3. If for instance

$$\begin{aligned}
\boldsymbol{e}_2 &= \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, 0, \ldots, 0\right), \\
\boldsymbol{e}_3 &= \left(\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}}, 0, \ldots, 0\right), \\
&\;\;\vdots \\
\boldsymbol{e}_k &= ((k^2-k)^{-1/2}, \ldots, (k^2-k)^{-1/2}, -(k-1)(k^2-k)^{-1/2}),
\end{aligned}$$

then

$$\frac{U_n^2}{\sqrt{n}}, \frac{U_n^3}{\sqrt{n}}, \ldots, \frac{U_n^k}{\sqrt{n}}$$

converge in distribution to

$$\frac{1}{\sqrt{2}}(Y_1 - Y_2), \frac{1}{\sqrt{6}}(Y_1 + Y_2 - 2Y_3), \ldots, (k^2-k)^{-1/2}(Y_1 + \cdots + Y_{k-1} - (k-1)Y_k)$$

in the notation of Theorem 3.

References

[1] R. Azencott, Grandes déviations et applications, in: Ecole d'Eté de Probabilités de Saint-Flour VIII, Lecture Notes in Math., Vol. 774 (Springer, Berlin, 1980) pp. 1-176.

[2] G. Ben Arous and M. Brunaud, Méthode de Laplace: Etude variationnelle des fluctuations de diffusions de type "champ moyen", Stochastics and Stochastics Reports 31 (1990) 79-144.

[3] E. Bolthausen, Laplace approximations for sums of independent random vectors, Part II: Degenerate maxima and manifolds of maxima, Probab. Theory Rel. Fields 76 (1987) 167-206.

[4] N.R. Chaganty and J. Sethuraman, Limit theorems in the area of large deviations for some dependent random variables, Ann. Probab. 15 (1987) 628-645.

[5] D.A. Dawson, Critical dynamics and fluctuations for a mean-field model of cooperative behaviour, J. Statist. Phys. 31 (1983) 29–85.

[6] F. Dunlop and C.M. Newman, Multicomponent field theories and classical rotators, Commun. Math. Phys. 44 (1975) 223–235.

[7] R.S. Ellis and C.M. Newman, Limit theorems for sums of dependent random variables occurring in statistical mechanics, Z. Wahrsch. Verw. Gebiete 44 (1978) 117–139.

[8] R.S. Ellis, Entropy, Large Deviations and Statistical Mechanics (Springer, New York, 1985).

[9] W. Feller, An Introduction to Probability Theory and its Applications, Vol. II (Wiley, New York, 1971, 2nd ed.).

[10] M. Kac, Mathematical mechanisms of phase transition, in: M. Chretien, E.P. Gross and S. Deser, eds., Statistical Physics: Phase Transitions and Superfluidity, Vol. 1, Brandeis University Summer Institute in Theoretical Physics (Gordon and Breach, New York, 1968) pp. 241–305.

[11] A. Martin-Löf, A Laplace approximation for sums of independent random variables, Z. Wahrsch. Verw. Gebiete 59 (1982) 101–115.

[12] F. Papangelou, On the Gaussian fluctuations of the critical Curie–Weiss model in statistical mechanics, Probab. Theory Rel. Fields 83 (1989) 265–278.

[13] V.V. Petrov, Sums of Independent Random Variables, Ergeb. Math. Grenzgeb., Vol. 82 (Springer, Berlin, 1975).

[14] B. Simon and R.B. Griffiths, The $(\varphi^4)_2$ field theory as a classical Ising model, Commun. Math. Phys. 33 (1973) 145–164.

[15] C.J. Thompson, Classical Equilibrium Statistical Mechanics (Clarendon, Oxford, 1988).